AD A093633

MRC Technical Summary Report #2119

SMOOTHING ESTIMATION OF
STOCHASTIC PROCESSES - PART I:
CHANGE OF INITIAL CONDITION FORMULAE

V. Solo

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

September 1980

(Received May 20, 1980)

DTIC FILE COPY

80 12 22 046

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

SMOOTHING ESTIMATION OF STOCHASTIC PROCESSES - PART I:
CHANGE OF INITIAL CONDITION FORMULAE
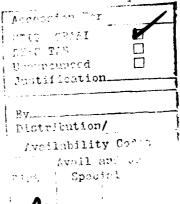
V. Solo

Technical Summary Report #2119
September 1980

ABSTRACT

The change of initial condition (CIC) problem in the theory of smoothing

in linear estimation is formulated in terms of a fixed rank perturbation to a

covariance kernel. The CIC or partitioning formulae are then shown to apply

for general nonstationary processes. It becomes clear that the formulae

involved derive from inversion formulae for fixed rank modification of a

positive definite kernel or matrix.

Accession For

DTIC GRA&I

DTIC TAB ☐

Unannounced ☐

Justification

By

Distribution/

Availability Codes

Avail and/or

Dist Special

A

SIGNIFICANCE AND EXPLANATION

The estimation of one process from measurements on another related

process is a problem that arises in many areas such as Time Series Analysis,

Econometrics, Communications Engineering and Control Engineering. Particularly

in the Engineering Applications there is a great interest in various computa-

tional forms of the algorithms proposed to solve the above problem. One

important question concerns the effect of perturbations, in the initial condi-

tions of the process statistics, on the algorithms. This article discusses

such effects for general process models and relates the results to the inver-

sion of perturbed matrices or kernels.

SMOOTHING ESTIMATION OF STOCHASTIC PROCESSES - PART I:
CHANGE OF INITIAL CONDITION FORMULAE

V. Solo

Introduction. Recently a great amount of attention has been focussed on various algorithms for solving the smoothing problem of linear estimation theory, see Ljung and Kailath [11], Lainiotis [10], and references in these. The main algorithms include the two-filter formulae, Mayne [13], Fraser [3], Mehra [14]; the innovations formulae, Kailath and Frost [7]; and the partitioned algorithms, Govindaraj and Lainiotis [4].

This work is the first part of a two part investigation of these algorithms. In Part I it is shown how change of initial condition (CIC) or partitioning formulae hold in a very general setting (the CIC problem is shown to involve fixed rank perturbation in matrix inversion). In Part II the nature of the two-filter algorithms is explored by providing a simple derivation that shows to what extent the formulae hold generally and so reveals exactly how a wide sense Markovian assumption in necessary for their full utility.

The remainder of the paper is structured as follows. Section I contains a discussion of CIC formulae for discrete observations. Section II concerns CIC formulae for continuous observations (actually the formulae are the same). Section III discusses the relation with other work.

Before continuing recall the matrix inversion lemma (MIL). Let $\tau, \ldots, \Gamma, D$ be matrices related by

$$Z = \Box + \Gamma D \Gamma'$$

then

$$Z^{-1} = (\Box + \Gamma D \Gamma')^{-1} = \Box^{-1} - \Box^{-1} \Gamma (D^{-1} + \Gamma' \Box^{-1} \Gamma)^{-1} \Gamma' \Box^{-1} \qquad \text{(MILa)}$$

The following consequences are well known

$$Z^{-1} = \Box^{-1} - Z^{-1} \Gamma D \Gamma' \Box^{-1} \qquad \text{(MILb)}$$

$$z^{-1}(I + \ldots z^{-1}) = \ldots^{-1} \qquad \text{(III.c)}$$

$$\ldots^{-1} = {}^{-1}(I - \ldots z^{-1}) \qquad \text{(III.d)}$$

The next two relations are less often used. By comparing (III.c), (III.d) it follows that

$$(I - \ldots^{-1})^{-1} = I + \ldots^{-1}$$

or

$$(I - \ldots^{-1})^{-1} = \ldots^{-1} + \ldots^{-1} \qquad \text{(III.e)}$$

Similarly set $\ldots$ in (III.d) to see

$$(\ldots + \ldots)^{-1} = {}^{-1} - (\ldots^{-1} + \ldots)^{-1} \qquad \text{(III.f)}$$

I. Change of Initial Condition Formulae.

Consider the linear estimation of an n_x dimensional process $\underline{x}(t)$ from measurements on a related n_y dimensional process $\underline{y}(t)$. We deal firstly with discrete measurements (at times $t_1 < t_2 < \ldots < t_N = T$ in an observation interval $0 \le t \le T$) which we collect into a vector $\underline{y}$. Suppose the process $\underline{x}(t)$ is composed of a zero mean process $\underline{x}_a(t)$ plus a transient component of the form $\underline{\iota}(t,0)\underline{\xi}$ where $\underline{\iota}(t,0)$ (henceforth denoted $\underline{\iota}_t$) is a deterministic transition matrix (with viz : $\underline{\iota}(0,0) = \underline{I}$) and $\underline{\xi}$ is a random vector uncorrelated with $\underline{x}_a(t)$. It is supposed that the transient has induced a perturbation within the measured process $\underline{y}(t)$ so that $\underline{y}(t)$ is of the form $\underline{y}(t) = \underline{y}_a(t) + \underline{g}(t,0)\underline{\xi}$ where $\underline{y}_a(t)$ is a zero mean process and $\underline{g}(t,0)$ a deterministic function. Thus the measurement vector has the form

$$\underline{y} = \underline{y}_a + \underline{G}\underline{\xi}$$

The setting described here has been used by Aasnaes and Kailath [1] to study the robustness of filtering algorithms to initial-condition perturbations.

Ia. A basic CIC Formula.

Consider now the linear estimation of $\underline{x}(t)$ from the measurements $\underline{y}$ under varying assumptions for the statistics of the $\underline{\xi}$ vector i.e. the initial conditions for $\underline{x}(t)$. Under condition "0" let $F_0(\underline{\xi})$ be the distribution function of the vector $\underline{\xi}$ and suppose

$$E_0(\underline{\xi}) = \underline{m}_0 \quad , \quad E_0(\underline{\xi} - \underline{m}_0)(\underline{\xi} - \underline{m}_0)' = \underline{\Pi}_0$$

where E_0 denotes integration with respect to the distribution $F_0(\underline{\xi})$. We have for example

$$E_0(\underline{x}(t)) = \underline{\iota}_t \underline{m}_0$$

Accordingly we can compute the linear least squares estimate

$$\underline{\hat{x}}_0(t|T) = E_0(\underline{x}(t)) + E_0(\underline{x}_0(t)\underline{W}_0')E_0(\underline{W}_0\underline{W}_0')^{-1}\underline{W}_0 \tag{1a}$$

where

$$\underline{x}_0(t) = \underline{x}(t) - E_0(\underline{x}(t)) = \underline{x}_a(t) + \underline{\lambda}_t(\underline{1} - \underline{m}_0) \tag{1b}$$

$$\underline{W}_0 = \underline{W} - E_0(\underline{W}) = \underline{W} - \underline{m}_0 = \underline{W}_a + \underline{\beta}(\underline{1} - \underline{m}_0) \tag{1c}$$

Now suppose the initial conditions are changed to "1"

$$E_1(\underline{1}) = \underline{m}_1 \quad , \quad E_1(\underline{1} - \underline{m}_1)(\underline{1} - \underline{m}_1)' = \underline{\pi}_{-1}$$

Thus we are interested in computing $\underline{\hat{x}}_1(t|T)$ from the condition "0" calculations without having to use the formula

$$\underline{\hat{x}}_1(t|T) = E_1(\underline{x}(t)) + E_1(\underline{x}_1(t)\underline{W}_1')E_1(\underline{W}_1\underline{W}_1')^{-1}\underline{W}_1 \tag{2a}$$

where

$$\underline{x}_1(t) = \underline{x}(t) - E_1(\underline{x}(t)) = \underline{x}_a(t) + \underline{\lambda}_t(\underline{1} - \underline{m}_1) \tag{2b}$$

$$\underline{W}_1 = \underline{W} - E_1(\underline{W}) = \underline{W} - \underline{m}_1 = \underline{W}_a + \underline{\beta}(\underline{1} - \underline{m}_1) \tag{2c}$$

There from (1b) and (2b) that

$$E_1(\underline{x}(t)) - E_0(\underline{x}(t)) = \underline{\lambda}_t(\underline{m}_1 - \underline{m}_0) \tag{3}$$

Now denote

$$\underline{M}_1 = E_1(\underline{x}_1(t)\underline{W}_1') \quad , \quad \underline{S}_1 = E_1(\underline{W}_1\underline{W}_1') \tag{4}$$

and similarly $\underline{M}_0$, $\underline{S}_0$. Thus from (1a)

$$\underline{\hat{x}}_1(t|T) - E_1(\underline{x}(t)) = \underline{M}_1\underline{S}_1^{-1}\underline{W}_1 \tag{5a}$$

$$\underline{\hat{x}}_0(t|T) - E_0(\underline{x}(t)) = \underline{M}_0\underline{S}_0^{-1}\underline{W}_0 \tag{5b}$$

Consider that from (2b), (2c)

$$\underline{M}_1 = E_1(\underline{x}_1(t)\underline{W}_1')$$

$$= E_1[\underline{x}_a(t) + \underline{\lambda}_t(\underline{1} - \underline{m}_1)][\underline{W}_a' + (\underline{1} - \underline{m}_1)'\underline{\beta}']$$

$$= E(\underline{x}_a(t)\underline{W}_a') + \underline{\lambda}_t\underline{\pi}_{-1}\underline{\beta}' \tag{6a}$$

-4-

No subscript is given to this last expectation since the same result is obtained under "0" or "1" Similarly

$$\underline{M}_0 = E(\underline{x}_a(t)\underline{\mu}'_a) + \underline{\Delta}_t\underline{\Xi}_0\underline{G}' \tag{6b}$$

Thus

$$\underline{M}_1 = \underline{M}_0 + \underline{\Delta}_t\underline{\Delta}\underline{G}' \; ; \; \underline{\Delta} = \underline{\Xi}_1 - \underline{\Xi}_0 \tag{6c}$$

Similarly

$$\underline{S}_1 = \underline{S}_0 + \underline{G}\underline{\Delta}\underline{G}' \tag{6d}$$

Thus $\underline{M}_1$, $\underline{S}_1$, are perturbations of $\underline{M}_0$, $\underline{S}_0$. When further observations are added $\underline{S}_1$ remains a fixed rank perturbation of $\underline{S}_0$. This point is the essence of the ensuing formulae.

Returning to (5a) we have

$$\hat{\underline{x}}_1(t|T) - E_1(\underline{x}(t)) = \underline{M}_1\underline{S}_1^{-1}\underline{\mu}_1$$

$$= \underline{M}_1\underline{S}_1^{-1}\underline{\mu}_0 - \underline{M}_1\underline{S}_1^{-1}\underline{G}(\underline{m}_1 - \underline{m}_0)$$

$$= \underline{M}_1(\underline{S}_0^{-1} - \underline{S}_1^{-1}\underline{G}\underline{\Delta}\underline{G}'\underline{S}_0^{-1})\underline{\mu}_0 - \underline{M}_1\underline{S}_1^{-1}\underline{G}(\underline{m}_1 - \underline{m}_0) \quad \text{by (6d) and (MILE)}$$

$$= \underline{M}_1\underline{S}_0^{-1}\underline{\mu}_0 - \underline{M}_1\underline{S}_1^{-1}\underline{G}\underline{\Delta}\underline{G}'\underline{S}_0^{-1}\underline{\mu}_0 - \underline{M}_1\underline{S}_1^{-1}\underline{G}(\underline{m}_1 - \underline{m}_0)$$

$$= \underline{M}_0\underline{S}_0^{-1}\underline{\mu}_0 + \underline{\Delta}_t\underline{\Delta}\underline{G}'\underline{S}_0^{-1}\underline{\mu}_0 - \underline{M}_1\underline{S}_1^{-1}\underline{G}\underline{\Delta}\underline{G}'\underline{S}_0^{-1}\underline{\mu}_0 - \underline{M}_1\underline{S}_1^{-1}\underline{G}(\underline{m}_1 - \underline{m}_0) \quad \text{by (6c)}$$

$$= \hat{\underline{x}}_0(t|T) - E_0(\underline{x}(t)) + (\underline{\Delta}_t - \underline{M}_1\underline{S}_1^{-1}\underline{G})\underline{\Delta}\underline{G}'\underline{S}_0^{-1}\underline{\mu}_0 - \underline{M}_1\underline{S}_1^{-1}\underline{G}(\underline{m}_1 - \underline{m}_0) \quad \text{by (5a)}$$

$$= \hat{\underline{x}}_0(t|T) - E_0(\underline{x}(t)) + (\underline{\Delta}_t - \underline{M}_1\underline{S}_1^{-1}\underline{G})(\underline{\Delta}\underline{G}'\underline{S}_0^{-1}\underline{\mu}_0 + \underline{m}_1 - \underline{m}_0) - \underline{\Delta}_t(\underline{m}_1 - \underline{m}_0)$$

or, recalling (3) we find

$$\hat{\underline{x}}_1(t|T) = \hat{\underline{x}}_0(t|T) + (\underline{\Delta}_t - \underline{M}_1\underline{S}_1^{-1}\underline{G})(\underline{\Delta}\underline{G}'\underline{S}_0^{-1}\underline{\mu}_0 + \underline{m}_1 - \underline{m}_0) \tag{6e}$$

Let us write

$$\underline{m}_0(0|T) = \underline{\Delta}'\underline{S}_0^{-1}\underline{\mu}_0 \tag{7a}$$

-5-

and derive that

$$\underline{x}_0(0,T) = E_0(\underline{x}(0)) + E_0(\underline{x}_a(0)\underline{w}'_0)E_0(\underline{w}_0\underline{w}'_0)^{-1}\underline{w}_0$$

$$= \underline{m}_0 + E_0(\underline{x}_a(0) - \underline{1} - \underline{m}_0)(\underline{w}'_a + (\underline{1} - \underline{m}_0)'\underline{1}')\underline{w}_0^{-1}\underline{w}_0$$

$$= \underline{m}_0 + \underline{I}_0\underline{1}'\underline{S}_0^{-1}\underline{w}_0 \quad \text{since} \quad E_a(0) = \underline{1} \ .$$

$$= \underline{m}_0 + \underline{I}_0\underline{I}_0(0,T) \tag{7i}$$

and we can write

$$\underline{x}_1(0,T) = \underline{x}_0(t,T) + (\underline{I}_t - \underline{H}_1\underline{S}_1^{-1}\underline{S})(\underline{m}_1 - \underline{m}_0 + \underline{I}_0(0,T))$$

$$= \underline{x}_0(t,T) + (\underline{I}_t - \underline{H}_1\underline{S}_1^{-1}\underline{S})(\underline{m}_1 - \underline{m}_0 + \underline{I}_0^{-1}(\underline{x}_0(0,T) - \underline{m}_0))$$

Finally introduce for $t \geq s$

$$\underline{I}_1(t,s,T) = E_1\{[\underline{x}(t) - \underline{x}_1(t,T)][\underline{x}'(s) - \underline{x}'_1(s,T)]\}$$

$$= E_1\{(\underline{x}(t) - \underline{x}_1(t,T))\underline{x}'(s)\}$$

because the least squares estimation error is orthogonal to any linear function of the measurements.

Then we find

$$\underline{I}_1(t,s,T) = E_1(\underline{x}(t) - \underline{x}_1(t,T))\underline{x}'(s)$$

$$= E_1(\underline{x}_1(t) + \underline{1}_t(\underline{1} - \underline{m}_1) - \underline{H}_1\underline{S}_1^{-1}\underline{w}_1)(\underline{x}'_a(s) + \underline{1}')$$

$$= \underline{I}_{t-1} - \underline{H}_1\underline{S}_1^{-1}\underline{H}_1 \quad \text{since} \quad \underline{x}_a(0) = \underline{1} \tag{7c}$$

$$\underline{x}_1(0,T) = \underline{x}_0(t,T) + \underline{I}_1(0,s,T)\underline{I}_1^{-1}(\underline{m}_1 - \underline{m}_0 + \underline{I}_0(0,T)) \tag{8a}$$

$$= \underline{x}_0(t,T) + \underline{I}_1(t,s,T)\underline{I}_1^{-1}(\underline{m}_1 - \underline{m}_0 + \underline{I}_0^{-1}(\underline{x}_0(0,T) - \underline{m}_0)) \tag{8b}$$

These equations have been introduced to allow a formula that holds when $\underline{I}_0 = \underline{0}$.

$$\underline{x}_1(0,T) = \underline{I}_0(t,s,T)\underline{I}_1^{-1} \ldots$$

$$\underline{\Psi}_1(t,0|T) = \underline{P}_1(t,0|T)\underline{\Pi}_1^{-1} = \underline{i}_t - \underline{M}_1\underline{S}_1^{-1}\underline{G}$$

Note that in formulae (8) the suffices 0,1 are dummy suffices and may be interchanged. Since no relation has been specified between t and T, equations (8) are both filtering and smoothing formulae. It is interesting that the data dependent portion of the correction term in (8) is the same for all t.

Now let us turn to find a CIC formula for $\underline{P}_1(t,0|T)$ or $\underline{\Psi}_1(t,0|T)$. Since $\underline{\lambda}_1(0|T)$ is of intermediate interest a CIC for this will be found too.

Ib. <u>A CIC formula for</u> $\underline{P}_1(t,0|T)\underline{\Pi}_1^{-1}$.

Consider then

$$\underline{P}_1(t,0|T)\underline{\Pi}_1^{-1} = \underline{i}_t - \underline{M}_1\underline{S}_1^{-1}\underline{G}$$

$$= \underline{i}_t - \underline{M}_0\underline{S}_1^{-1}\underline{G} - \underline{i}_t\underline{\Delta}\underline{G}'\underline{S}_1^{-1}\underline{G} \quad \text{by} \quad (6c)$$

$$= \underline{i}_t(\underline{I} - \underline{\Delta}\underline{G}'\underline{S}_1^{-1}\underline{G}) - \underline{M}_0\underline{S}_1^{-1}\underline{G}$$

$$= \underline{i}_t(\underline{I} - \underline{\Delta}\underline{G}'\underline{S}_1^{-1}\underline{G}) - \underline{M}_0\underline{S}_0^{-1}\underline{G}(\underline{I} - \underline{\Delta}\underline{G}'\underline{S}_1^{-1}\underline{G}) \quad \text{by} \quad (6d) \quad \text{and} \quad (\text{MILd})$$

$$= (\underline{i}_t - \underline{M}_0\underline{S}_0^{-1}\underline{G})(\underline{I} - \underline{\Delta}\underline{G}'\underline{S}_1^{-1}\underline{G})$$

$$= \underline{P}_0(t,0|T)\underline{\Pi}_0^{-1}(\underline{I} - \underline{\Delta}\underline{G}'\underline{S}_1^{-1}\underline{G}) \tag{9a}$$

let us call

$$\underline{\Gamma} = \underline{\Delta} - \underline{\Delta}\underline{G}'\underline{S}_1^{-1}\underline{G}\underline{\Delta} = (\underline{I} - \underline{\Delta}\underline{G}'\underline{S}_1^{-1}\underline{G})\underline{\Delta} \tag{10a}$$

so that

$$\underline{P}_0(t,0|T)\underline{\Pi}_0^{-1} = \underline{\Psi}_0(t,0|T)$$

$$= \underline{P}_1(t,0|T)\underline{\Pi}_1^{-1}\underline{\Gamma}^{-1} = \underline{\Psi}_1(t,0|T)\underline{\Delta}\underline{\Gamma}^{-1} \tag{9b}$$

On the other hand, guided by (MILe) we find via (6d) that

-7-

$$\underline{\Sigma}^{-1} = (\underline{\Sigma} - \underline{M}_0'\underline{S}_1^{-1}\underline{M}_0)^{-1} = \underline{\Sigma}^{-1} + \underline{M}'\underline{S}_0^{-1}\underline{M} \tag{10a}$$

$$= \underline{\Sigma}^{-1} + \underline{Q}_0 \tag{10b}$$

where, recalling (7a) and (4)

$$\underline{Q}_0 = \underline{M}'\underline{S}_0^{-1}\underline{M} = E_0(\underline{z}_0(0|T)\,\underline{z}_0(0|T)') \tag{10c}$$

Recalling (7b) we can view $\underline{Q}_0$ in another way by introducing

$$\underline{L}(0,0|T) = E_0(\hat{\underline{x}}_0(0|T) - \underline{m}_0)(\hat{\underline{x}}_0(0|T) - \underline{m}_0)'$$

$$= \underline{\Sigma}_0 E_0(\underline{z}_0(0|T)\,\underline{z}_0(0|T)')\underline{\Sigma}_0$$

$$= \underline{\Sigma}_0 \underline{Q}_0 \underline{\Sigma}_0' \tag{10d}$$

Using (10c) we can write (9a) as

$$\underline{L}_1(t,0|T) = \underline{L}_0(t,0|T)\underline{\Sigma}^{-1}$$

$$= \underline{L}_0(t,0|T)(\underline{\Sigma}^{-1} + \underline{Q}_0)^{-1}\underline{\Sigma}^{-1} \tag{9c}$$

$$= \underline{L}_0(t,0|T)(\underline{I} - \underline{Q}_0\underline{\Sigma})$$

$$= \underline{L}_0(t,0|T) - \underline{L}_0(t,0|T)\underline{Q}_0\underline{\Sigma} \tag{9d}$$

There is a nice formula for $\underline{Q}_1$. From (10a)

$$\underline{\Sigma}^{-1}\underline{\Sigma}^{-1} = \underline{\Sigma}^{-1} - \underline{Q}_1$$

so that by (X11f)

$$\underline{Q}_1^{-1} = \underline{\Sigma} - (\underline{\Sigma}^{-1} - \underline{\Sigma}^{-1})^{-1}$$

$$= \underline{\Sigma} + \underline{Q}_0^{-1} \quad \text{by (10c)} \tag{11f}$$

Given $\underline{\Sigma}$, we can use (10c) or (10f), (10a) to find $\underline{\Sigma}$.

-8-

Ic. $\underline{\text{CIC formula for}}$ $\underline{\lambda}_1(0|T)$.

Now

$$\underline{\lambda}_1(0|T) = \underline{G}'\underline{S}_1^{-1}\underline{\gamma}_1$$

$$= \underline{G}'\underline{S}_1^{-1}\underline{\mu}_0 - \underline{G}'\underline{S}_1^{-1}\underline{G}(\underline{m}_1 - \underline{m}_0)$$

$$= \underline{G}'(\underline{S}_0^{-1} - \underline{S}_1^{-1}\underline{G}\underline{\Lambda}\underline{G}'\underline{S}_0^{-1})\underline{\mu}_0 - \underline{Q}_1(\underline{m}_1 - \underline{m}_0) \quad \text{by (MILb) and (10d)}$$

$$= \underline{\lambda}_0(0|T) - \underline{G}'\underline{S}_1^{-1}\underline{G}\underline{\Lambda}\underline{\lambda}_0(0|T) - \underline{Q}_1(\underline{m}_1 - \underline{m}_0)$$

$$= (\underline{I} - \underline{G}'\underline{S}_1^{-1}\underline{G}\underline{\Lambda})\underline{\lambda}_0(0|T) - \underline{Q}_1(\underline{m}_1 - \underline{m}_0)$$

$$= \underline{\Lambda}^{-1}\underline{\Lambda}\underline{\lambda}_0(0|T) - \underline{Q}_1(\underline{m}_1 - \underline{m}_0) \quad \text{by (10a)}$$

$$= \underline{\Lambda}^{-1}\underline{\Lambda}(\underline{\lambda}_0(0|T) - \underline{\Lambda}^{-1}\underline{\Lambda}\underline{Q}_1(\underline{m}_1 - \underline{m}_0))$$

However by (10c)

$$\underline{\Lambda}^{-1}\underline{\Lambda}\underline{Q}_1 = (\underline{I} + \underline{Q}_0\underline{\Lambda})\underline{Q}_1 = \underline{Q}_0(\underline{Q}_0^{-1} + \underline{\Lambda})\underline{Q}_1$$

$$= \underline{Q}_0 \quad \text{by (10f)}$$

Thus

$$\underline{\lambda}_1(0|T) = \underline{\Lambda}^{-1}\underline{\Lambda}(\underline{\lambda}_0(0|T) - \underline{Q}_0(\underline{m}_1 - \underline{m}_0)) \tag{11a}$$

Also observe that

$$\underline{m}_1 - \underline{m}_0 + \underline{\lambda}_1(0|T) = \underline{\lambda}_0(0|T) + (\underline{I} - \underline{\Lambda}\underline{Q}_0)(\underline{m}_1 - \underline{m}_0)$$

$$= \underline{\Lambda}\underline{\Lambda}^{-1}(\underline{\lambda}_0(0|T) + \underline{m}_1 - \underline{m}_0) \quad \text{by (10a)} \tag{11b}$$

Id. $\underline{\text{A CIC formula for}}$ $\underline{P}_1(t,t|T)$.

In Appendix A it is shown that

$$\underline{P}_1(t,t|T) = \underline{P}_0(t,t|T) + \underline{P}_1(t,0|T)\underline{\Lambda}_1^{-1}\underline{\Lambda}_0^{-1}\underline{P}_0(t,0|T)$$

$$= \underline{P}_0(t,t|T) + \underline{P}_1(t,0|T)\underline{\Lambda}_0'(t,0|T) \tag{12a}$$

-9-

Application of (8b) allows this to be written

$$\underline{F}_1(t,t'|T) = \underline{F}_0(t,t'|T) + \underline{F}_0(t,0|T)\underline{\hat{x}}_1'(t,0|T) \tag{12b}$$

Finally note from orthogonality

$$\underline{P}_1 = \underline{P}_1(0,0|T) = \underline{\Sigma}_1 - \hat{\underline{P}}_1(0,0|T)$$

$$= \underline{\Sigma}_1 - \underline{\Sigma}_1 \underline{Q}_1 \underline{\Sigma}_1 \quad \text{by (10e)} \tag{12c}$$

Thus by (IIL.f)

$$\underline{P}_1^{-1} = \underline{\Sigma}_1^{-1} + (\underline{Q}_1^{-1} - \underline{\Sigma}_1)^{-1}$$

Similarly

$$\underline{P}_0^{-1} = \underline{\Sigma}_0^{-1} + (\underline{Q}_0^{-1} - \underline{\Sigma}_0)^{-1} \; ; \; \underline{P}_0 = \underline{P}_0(0,0|T)$$

now since $\underline{\Sigma}_1 - \underline{\Sigma}_0 = \underline{I}$, (10f) implies by subtraction

$$\underline{P}_1^{-1} - \underline{\Sigma}_1^{-1} = \underline{P}_0^{-1} - \underline{\Sigma}_0^{-1} \tag{10g}$$

Ic.

The full set of change of initial condition formulae are

Equations (8a) , (8b) for $\underline{\hat{x}}_1(t|T)$

Equations (9b) , (9c) for $\underline{\hat{P}}_1(t,0|T)$

Equations (12a) , (12b) for $\underline{P}_1(t,t|T)$

Equation (10e) for $\underline{\Sigma}$

Equation (10f) for $\underline{Q}_1$

Notice that every quantity in these equations has a stochastic interpretation. This makes for a rapid comparison with other work; see Section III. Also it is shown in Section II that these equations hold also for continuous time measurement.

If. Calculation of $\underline{\xi}(t,0|T)$ in continuous measurement State Space Models.

The quantity $\underline{\xi}(t,0|T) = \underline{P}(t,0|T)\underline{\pi}^{-1}$ has only appeared in earlier work with $T = t$ or $t = 0$. Here we derive a differential equation for $\underline{\xi}(t,0|T)$. Begin with the usual State Space Model

$$d\underline{x}/dt = \underline{F}(t)\underline{x}(t) + \underline{\Gamma}(t)\underline{w}(t)$$

$$\underline{y}(t) = \underline{H}(t)\underline{x}(t) + \underline{v}(t)$$

where $\underline{w},\underline{v}$ are uncorrelated white noises with covariance matrix $\underline{I}$. We use a formula of Rauch et. al., [21], (also Kailath and Frost [7, equation (34a)]

$$d\hat{\underline{x}}(t|T)/dt = \underline{F}(t)\hat{\underline{x}}(t|T) + \underline{\Gamma}(t)\underline{\Gamma}'(t)\underline{P}^{-1}(t)(\hat{\underline{x}}(t|T) - \hat{\underline{x}}(t|t))$$

Thus

$$d\underline{\xi}(t,0|T)/dt = d/dt \, E(\underline{x}(t) - \hat{\underline{x}}(t|T))\underline{x}'(0)\underline{\pi}^{-1}$$

$$= \underline{F}(t)E(\underline{x}(t) - \hat{\underline{x}}(t|T))\underline{x}'(0)\underline{\pi}^{-1}$$

$$- \underline{\Gamma}(t)\underline{\Gamma}'(t)\underline{P}^{-1}(t)[\underline{\xi}(t,0|t) - \underline{\xi}(t,0|T)]$$

Thus

$$d\underline{\xi}(t,0|T)/dt = (\underline{F}(t) + \underline{\Gamma}(t)\underline{\Gamma}'(t)\underline{P}^{-1}(t))\underline{\xi}(t,0|T)$$

$$= \underline{\Gamma}(t)\underline{\Gamma}'(t)\underline{P}^{-1}(t)\underline{\xi}(t,0|t)$$

We need only augment this with an equation for $\underline{\xi}(t,0|t)$ which is easily shown to be the Kalman Filter transition matrix

$$d\underline{\xi}(t,0|t)/dt = (\underline{F}(t) - \underline{P}(t)\underline{H}'(t)\underline{H}(t))\underline{\xi}(t,0|t)$$

with initial condition $\underline{\xi}(0,0|0) = \underline{I}$.

Ig. CIC formula for a Gaussian likelihood.

Consider the log Gaussian likelihood which is, to within a constant

$$\ln L_1 = \ln |\underline{S}_1| - \frac{1}{2}\underline{\mu}'_1\underline{S}_1^{-1}\underline{\mu}_1 \tag{13a}$$

Now

$$|\underline{S}_1| = |\underline{S}_0 + \underline{G}'\underline{\Delta}\underline{G}| = |\underline{S}_0||\underline{I} + \underline{S}_0^{-1}\underline{G}'\underline{\Delta}\underline{G}|$$

$$= |\underline{S}_0||\underline{I} + \underline{\Delta}\underline{G}'\underline{S}_0^{-1}\underline{G}|$$

$$= |\underline{S}_0||\underline{\Delta}||\underline{\Delta}^{-1} + \underline{Q}_0|$$

$$= |\underline{S}_0||\underline{\Delta}||\underline{\Sigma}^{-1}| \qquad \text{by (10c)} \qquad (13\text{b})$$

Next

$$\underline{\mathcal{Y}}_1'\underline{S}_1^{-1}\underline{\mathcal{Y}}_1 = (\underline{\mathcal{Y}}_0 - \underline{G}(\underline{m}_1 - \underline{m}_0))'\underline{S}_1^{-1}\underline{\mathcal{Y}}_1$$

$$= \underline{\mathcal{Y}}_0'\underline{S}_1^{-1}\underline{\mathcal{Y}}_1 - (\underline{m}_1 - \underline{m}_0)'\underline{\lambda}_1$$

where $\underline{\lambda}_1 = \underline{\lambda}_1(0|T)$; see (7a).

So

$$\underline{\mathcal{Y}}_1'\underline{S}_1^{-1}\underline{\mathcal{Y}}_1 = \underline{\mathcal{Y}}_1'(\underline{S}_0^{-1} - \underline{S}_1^{-1}\underline{G}\underline{\Delta}\underline{G}'\underline{S}_0^{-1})\underline{\mathcal{Y}}_0 - (\underline{m}_1 - \underline{m}_0)'\underline{\lambda}_1$$

by (MILb). Continuing

$$\underline{\mathcal{Y}}_1'\underline{S}_1^{-1}\underline{\mathcal{Y}}_1 = \underline{\mathcal{Y}}_1'\underline{S}_0^{-1}\underline{\mathcal{Y}}_0 - \underline{\lambda}_1'\underline{\Delta}\underline{\lambda}_0 - (\underline{m}_1 - \underline{m}_0)'\underline{\lambda}_1$$

$$= \underline{\mathcal{Y}}_0'\underline{S}_0^{-1}\underline{\mathcal{Y}}_0 - (\underline{m}_1 - \underline{m}_0)'\underline{\lambda}_0 - \underline{\lambda}_1'\underline{\Delta}\underline{\lambda}_0 - (\underline{m}_1 - \underline{m}_0)'\underline{\lambda}_1$$

since $\underline{\mathcal{Y}}_1 = \underline{\mathcal{Y}}_0 - \underline{G}(\underline{m}_1 - \underline{m}_0)$.

Thus

$$\underline{\mathcal{Y}}_1^{-1}\underline{S}_1^{-1}\underline{\mathcal{Y}}_1 = \underline{\mathcal{Y}}_0'\underline{S}_0^{-1}\underline{\mathcal{Y}}_0 - [\underline{\Delta}^{-1}(\underline{m}_1 - \underline{m}_0) + \underline{\lambda}_0]'[(\underline{m}_1 - \underline{m}_0) + \underline{\Delta}\underline{\lambda}_1]$$

$$+ (\underline{m}_1 - \underline{m}_0)'\underline{\Delta}^{-1}(\underline{m}_1 - \underline{m}_0)$$

Now denote

$$\underline{\hat{x}}_1 = \underline{\hat{x}}_1(0|T) \; ; \; \underline{\hat{x}}_0 = \underline{\hat{x}}_0(0|T)$$

$$\underline{P}_1 = \underline{P}_1(0,0|T) \; ; \; \underline{P}_0 = \underline{P}_0(0,0|T)$$

and set $t = 0$ in (8a) to see

$$\Xi_1 P_1^{-1}(\hat{\underline{x}}_1 - \hat{\underline{x}}_0) = \underline{m}_1 - \underline{m}_0 + \dot{\underline{\lambda}}_0$$

also applying the inverse of (9b) with $t = 0$ gives

$$\Xi^{-1} \Xi_1 P_1^{-1} = \Xi_0 P_0^{-1}$$

so

$$\Xi_0 P_0^{-1}(\hat{\underline{x}}_1 - \hat{\underline{x}}_0) = \Xi^{-1}(\underline{m}_1 - \underline{m}_0 + \dot{\underline{\lambda}}_0)$$

$$= \Xi^{-1}(\underline{m}_1 - \underline{m}_0 + \dot{\underline{\lambda}}_1) \quad \text{by (11b)}$$

Thus

$$\underline{\mu}_1' S_1^{-1} \underline{\mu}_1 = \underline{\mu}_0' S_0^{-1} \underline{\mu}_0 - (\hat{\underline{x}}_1 - \hat{\underline{x}}_0)' P_1^{-1} \Xi_1 \Xi_0 P_0^{-1}(\hat{\underline{x}}_1 - \hat{\underline{x}}_0)$$

$$+ (\underline{m}_1 - \underline{m}_0)' \Xi^{-1}(\underline{m}_1 - \underline{m}_0)$$

However (12a) with $t = 0$ implies

$$\underline{P}_1 - \underline{P}_0 = \underline{P}_1 \Xi_1^{-1} \Xi_0^{-1} \underline{P}_0$$

Thus

$$\underline{\mu}_1' S_1^{-1} \underline{\mu}_1 = \underline{\mu}_0' S_0^{-1} \underline{\mu}_0 - (\hat{\underline{x}}_1 - \hat{\underline{x}}_0)'(\underline{P}_1 - \underline{P}_0)^{-1}(\hat{\underline{x}}_1 - \hat{\underline{x}}_0)$$

$$+ (\underline{m}_1 - \underline{m}_0)' \Xi^{-1}(\underline{m}_1 - \underline{m}_0) \qquad (13c)$$

So finally collecting (13a), (13b), (13c)

$$\ln L_1 = \ln L_0 + \ln |\Xi_1| - \ln |\Xi_0|$$

$$+ \frac{1}{2}(\hat{\underline{x}}_1(0|T) - \hat{\underline{x}}_0(0|T))'(\underline{P}_1(0,0|T) - \underline{P}_0(0,0|T))^{-1}(\hat{\underline{x}}_1(0|T) - \hat{\underline{x}}_0(0|T))$$

$$- \frac{1}{2}(\underline{m}_1 - \underline{m}_0)' \Xi^{-1}(\underline{m}_1 - \underline{m}_0) \qquad (14)$$

II. Continuous Time.

Let us retain the assumptions of Section I concerning the presence of the transient term in both $\underline{x}$ and $\underline{y}$ and suppose now that $\underline{y}$ is measured continuously over the interval $0 \leq t \leq T$. Introduce similarly to before the kernels M_i, S_i, $g \underline{\Delta} g'$, $\underline{\Delta}_t \underline{\Delta} g'$ defined on $[0,T] \times [0,T]$ by

$$M_i(t,s) = E_i(\underline{x}_i(t)\underline{y}_i(s)) \qquad i = 0, 1 \tag{15a}$$

$$S_i(t,s) = E_i(\underline{y}_i(t)\underline{y}_i'(s)) \qquad i = 0, 1 \tag{15b}$$

$$g \cdot g'(t,s) = \underline{g}(t)\underline{\Delta}g'(s) \ ; \ \underline{\Delta}_t \underline{\Delta}g'(t,s) = \underline{\Delta}_t\underline{\Delta}g'(s) \tag{15c}$$

where

$$\underline{x}_i(t) = \underline{x}(t) - E_i(\underline{x}(t)) \qquad i = 0, 1$$

$$\underline{y}_i(t) = \underline{y}(t) - E_i(\underline{y}(t)) \qquad i = 0, 1$$

Then it follows as in Section Ia that

$$S_1 = S_0 + g \underline{\Delta} g' \tag{15d}$$

$$M_1 = M_0 + \underline{\Delta} \underline{\Delta} g' \tag{15e}$$

suppose too, without loss of generality that

$$\underline{\Delta} \text{ is positive definite} \tag{15f}$$

Recall $\underline{g}(t)$ is an $n_y \times n_x$ matrix function. In the sequel $\underline{h}(t)$, $\underline{f}(t)$ will denote general n_y-vector functions.

The appropriate setting for the further development of CIC formulae in a general fashion is the theory of Reproducing Kernel Hilbert Spaces (RKHS): See Parzen [16] and Kailath [5]. The idea in the RKHS approach is to provide a coordinate free representation of continuous time equivalents of discrete forms such as "$\underline{h}'S_0^{-1}\underline{f}$."

Now Parzen [16] only discusses RKHS's for scalar processes so it will be indicated briefly how to define an appropriate RKHS for vector processes.

-14-

Proposition 1. Existence and Uniqueness of a vector RKHS generated by an $n_y \times n_y$ matrix covariance kernel (cf. Parzen [17, Theorem 3.1]). Let $\underline{y}(t)$, t in $[0,T]$ have covariance kernel $\underline{\Sigma}(t,s)$. Let $H(\Sigma)$ consist of all n_y-vector functions $\underline{h}(\cdot)$ on $[0,T]$ of the form

$$\underline{h}(t) = E(\underline{y}(t)U)$$

for some scalar U in $SL_2(\underline{y}(t)$, t in $[0,T])$ where SL_2 is the (scalar) Hilbert Space of random variables that are finite scalar linear combinations (of the form $\sum_i \underline{a}_i' \underline{y}(t_i)$) of random variables $\underline{y}(t)$, t in $[0,T]$ or limits in mean square of such finite linear combinations. Define an inner product on $H(\Sigma)$ as follows. If $\underline{h}(t) = E(\underline{y}(t)U)$; $\underline{f}(t) = E(\underline{y}(t)V)$ for $U, V \in SL_2$, define

$$\langle \underline{h}', \underline{f} \rangle_{\Sigma} = E(UV)$$

Then $H(\Sigma)$ is a Hilbert Space satisfying

(a) $\underline{\Sigma}_i(\cdot, t) \in H(\Sigma)$ $i = 1, \ldots, n_y$

where $\underline{\Sigma}_i(\cdot, t)$ is the ith column of $\underline{\Sigma}(\cdot, t)$

(b) $\langle \underline{h}', \underline{\Sigma}_i(\cdot, t) \rangle_{\Sigma} = h_i(t)$ for all $\underline{h} \in H(\Sigma)$

Proof. Follows much as in the proof of Theorem 3.1 of Parzen [17]. Observe for example though that

$$\underline{\Sigma}_i(s, t) = E(\underline{y}(s)y_i(t))$$

where $y_i(t)$ is the ith component of $\underline{y}(t)$. Thus

$$\langle \underline{f}', \underline{\Sigma}_i(\cdot, t) \rangle = E(y_i(t)V) = f_i(t)$$

Remark 1. Let $\underline{F}(t)$ be an $n_y \times m$ matrix function on the form $\underline{F}(t) = (\underline{f}_1(t) \ldots \underline{f}_m(t))$ where $\underline{f}_i(t)$ are n_y-vector functions. Also let $\underline{G}(t)$ be a similarly defined $n_y \times n$ matrix function. With a slight abuse of notation denote by $\langle \underline{F}, \underline{G} \rangle_{\Sigma}$ the $m \times n$ matrix whose i,j element is $\langle \underline{f}_i', \underline{g}_j \rangle$. Thus

$$\langle \underline{h}', \underline{\Sigma}_i(\cdot, t) \rangle = h_j(t)$$

-15-

can be written

$$\langle \underline{h}', \underline{\Sigma}(\cdot,t) \rangle_{\Sigma} = \underline{h}'(t)$$

Also

$$\langle \underline{\Sigma}_i'(\cdot,t), \underline{\Sigma}_j(\cdot,s) \rangle_{\Sigma} = \Sigma_{ij}(t,s)$$

reads

$$\langle \underline{\Sigma}(\cdot,t), \underline{\Sigma}(\cdot,s) \rangle_{\Sigma} = \underline{\Sigma}(t,s)$$

Finally we denote

$$\underline{f}_i \in H(\Sigma) \qquad i = 1, \ldots, m$$

by
$$\underline{F} \in H(\Sigma)$$

Remark 2. All the scalar RKHS theorems given by Parzen [16] apply to vector RKHS's.

The basic question in applying RKHS theory consists in finding computational expressions for the RKHS norm: for some concrete examples see Parzen [16] and Kailath et. al. [7]. To obtain results analagous to those of Section I we need a MIL for a RKHS. First however it is convenient to define the direct product of two vector RKHS's (cf. Parzen [17]; also Kailath [6, p. 18]).

If $\underline{Z}$, $\underline{\Sigma}$ are two m-dimensional positive definite kernels on $[0,T] \times [0,T]$ generating RKHS's $H(Z)$, $H(\Sigma)$ respectively then the direct product $H(Z) \otimes H(\Sigma)$ is the space of $m \times m$ matrix functions on $[0,T] \times [0,T]$ which are linear combinations of products of functions, one from $H(Z)$, the other from $H(\Sigma)$ (of the form $\underline{m}(t,s) = \sum_i \sum_j \underline{z}_i(t) \underline{\sigma}_j'(s) d_{ij})$ on limits in the norm

$$\| \underline{m}(\cdot,x) \|_{Z\Sigma}^2 = \sum_i \sum_j \langle \underline{z}_i', \underline{z}_j \rangle_Z \langle \underline{\sigma}_i', \underline{\sigma}_j \rangle_{\Sigma} d_{ij}$$

of such linear combinations (d_{ij} are real scalars).

Proposition 2. MIL for RKHS's. Let $\underline{Z}$, $\underline{\Sigma}$ be $m \times m$ matrix kernels on $[0,T] \times [0,T]$ related by

$$\underline{Z}(t,s) = \underline{\Sigma}(t,s) + \underline{\Sigma}(t)\underline{D}\underline{\Sigma}'(s)$$

where $\underline{\phi}(t)$ is an $m \cdot d$ matrix function and $\underline{D}$ a $d \cdot d$ positive definite matrix. Let $H(Z)$, $H(\phi)$ be the RKHS's generated by Z, ϕ respectively then iff $\underline{\phi} \in H(\phi)$ (ie. $\underline{\phi}_i \in H(\phi)$, $i = 1, \ldots, d$ where $\underline{\phi}_i(t)$ is the ith column of $\underline{\phi}(t)$) we have

    (a) $\underline{f} \in H(Z)$ iff $\underline{f} \in H(\phi)$

This does not mean $H(Z)$ is a subspace of $H(\phi)$ or vice versa since that would entail equality of inner products.

    (b) For $\underline{h}$, $\underline{f} \in H(Z)$ the $H(Z)$ inner product is

$$\langle \underline{h}', \underline{f} \rangle_Z = \langle \underline{h}', \underline{f} \rangle_\phi - \langle \underline{h}', \underline{\phi} \rangle V^{-1} \langle \underline{\phi}', \underline{f} \rangle$$

where $\underline{V} = \underline{D}^{-1} + \langle \underline{\phi}', \underline{\phi} \rangle$

Remark 1. Formula (b) was suggested by (MILa).

Remark 2. Note that $\underline{V}$ is well defined iff $\underline{\phi} \in H(\phi)$ i.e. $\langle \underline{\phi}', \underline{\phi} \rangle < \infty$.

Proof. To see (a) we use a theorem of Aronszajn [1, p. 354] which implies that (a) holds if

    (i) There are constants c_Z, c_ϕ such that $c_Z Z - \phi$ and $c_\phi \phi - Z$ are positive semidefinite.

In fact (i) is true with $c_Z = 1$ and

$$c_\phi = \lambda_D \, \mathrm{tr}(\langle \underline{\phi}', \underline{\phi} \rangle) \tag{16}$$

where λ_D = largest eigenvalue of $\underline{D}$. This claim is proven in Appendix B.

Now in view of (a), (b) is established if it is shown that

    (ii) $\underline{Z}(\cdot, t) \in H(\phi)$ for all t in $[0,T]$

    (iii) for all vectors $\underline{h} \in H(\phi)$

$$\langle \underline{h}', \underline{Z}(\cdot, t) \rangle_Z = \underline{h}'(t)$$

Now (ii) is clearly true since $\underline{\phi}(\cdot, t) \in H(\phi)$, $\underline{\phi} \in H(\phi)$ and $\underline{Z}(\cdot, t) = \underline{\phi}(\cdot, t) + \underline{\phi}(\cdot)\underline{D}\underline{\phi}'(t)$.

To see (iii) let $\underline{h} \in H(\phi)$ and compute

$$\langle \underline{h}',\underline{Z}(\cdot,t)\rangle_Z = \langle \underline{h}',\underline{Z}(\cdot,t)\rangle_{\Sigma} - \langle \underline{h}',\underline{\Gamma}\rangle_{\Sigma}\underline{V}^{-1}\langle \underline{\Gamma}',\underline{Z}(\cdot,t)\rangle_{\Sigma}$$

$$= \langle \underline{h}',\underline{\Sigma}(\cdot,t)\rangle_{\Sigma} + \langle \underline{h}',\underline{\Gamma}(\cdot)\underline{D}\underline{\Gamma}'(t)\rangle_{\Sigma}$$

$$- \langle \underline{h}',\underline{\Gamma}\rangle_{\Sigma}\underline{V}^{-1}\langle \underline{\Gamma}',\underline{\Sigma}(\cdot,t)\rangle_{\Sigma} - \langle \underline{h}',\underline{\Gamma}\rangle_{\Sigma}\underline{V}^{-1}\langle \underline{\Gamma}',\underline{\Gamma}(\cdot)\underline{D}\underline{\Gamma}'(t)\rangle_{\Sigma}$$

$$= \underline{h}'(t) + \langle \underline{h}',\underline{\Gamma}\rangle_{\Sigma}[\underline{I} - \underline{V}^{-1}\underline{D}^{-1} - \underline{V}^{-1}\langle \underline{\Gamma}',\underline{\Gamma}\rangle_{\Sigma}]\underline{D}\underline{\Gamma}'(t)$$

$$= \underline{h}'(t)$$

Since, by definition of $\underline{V}$

$$\underline{V}^{-1}(\underline{D}^{-1} + \langle \underline{\Gamma}',\underline{\Gamma}\rangle_{\Sigma}) = \underline{I}$$

The result is thus proven.

Remark 3. Since $\underline{\Gamma} \in H(\Sigma)$ and

$$\underline{\Gamma}(t)\underline{D}\underline{\Gamma}'(s) = \sum_{i=1}^{d}\sum_{j=1}^{d}d_{ij}\lambda_i(t)\lambda_j'(s)$$

it follows that $\underline{\Gamma}\underline{D}\underline{\Gamma}' \in H(\Sigma) \otimes H(\Sigma)$ i.e. (iv) holds

(iv) $\underline{Z} - \underline{\Sigma} \in H(\Sigma) \otimes H(\Sigma)$

Now however according to a result of Codaira [15] (see also Kailath [6, p. 20]) (i), (iv) are necessary and sufficient for the problem of "detection" between $\underline{Z}$ and $\underline{\Sigma}$ to be nonsingular. It is apparent that $\underline{\Gamma} \in H(\Sigma)$ ensures this nonsingularity.

Remark 4. In a similar way other MIL expressions analogous to those of Section I (viz: MIL(b) - MIL(f)) can be obtained (iff $\underline{\Gamma} \in H(\Sigma)$).

Now these results can be applied to the present problem.

Proposition 3. Let H_i be the RKHS's generated by $S_i(t,s)$, $i = 0, 1$. Suppose that $\underline{\Gamma}$ is positive definite. Then iff $\underline{g} \in H_0$ we have

(a) $\underline{h} \in H_0$ iff $\underline{h} \in H_1$

(b) The H_1 inner product is given by

$$\langle \underline{h}',\underline{f}\rangle_1 = \langle \underline{h}',\underline{f}\rangle_0 - \langle \underline{h}',\underline{g}\rangle_0\underline{\Gamma}^{-1}\langle \underline{g}',\underline{f}\rangle_0 \tag{17a}$$

where

$$\underline{\Gamma} = \underline{\Sigma}^{-1} + \langle \underline{g}',\underline{g}\rangle_0$$

Proof. Since $\underline{\underline{c}}$ is assumed positive definite the result follows from Proposition 2.

Remark. For future use we state the result corresponding to (MILb). Under the conditions of Proposition 3

$$\langle \underline{h}',\underline{f}\rangle_1 = \langle \underline{h}',\underline{f}\rangle_0 - \langle \underline{h}',\underline{g}\rangle_1 \underline{\underline{A}} \langle \underline{g}',\underline{f}\rangle_0 \tag{17b}$$

Finally we turn to consider the linear estimation problem. The RKHS framework allows a general solution to this problem: namely (Parzen [20])

$$\hat{\underline{x}}_i(t|T) = E_i(\underline{x}(t)) + \underline{U}_i(t|T) \qquad i = 0, 1$$

where $\underline{U}_i(t|T)$ is the unique solution to the equation

$$E_i(\underline{U}_i(t|T)\underline{Y}_i'(\cdot)) = \underline{M}_i(t,\cdot) = E_i(\underline{x}_i(t)\underline{Y}_i'(\cdot)) \quad \text{for all } \cdot \text{ in } [0,T] \quad i = 0, 1 \tag{18a}$$

where
$$\underline{Y}_i(\cdot) = \underline{Y}(\cdot) - \underline{g}(\cdot)\underline{m}_i \qquad i = 0, 1$$

That is, $\underline{U}_{ji}(t|T)$ (the jth element of $\underline{U}_i(t|T)$ is the unique member of $SL_2^{(i)}$ that satisfies (18a). We can write symbolically

$$\underline{U}_i(t|T) \qquad \langle \underline{M}_i(t,\cdot),\underline{Y}_i\rangle_i \qquad i = 0, 1 \tag{18b}$$

The equality is symbolic because $\underline{Y}_i \notin H_i$ (Indeed $\langle \underline{Y}_i',\underline{Y}_i\rangle_i$ has infinite variance-- this is clear since the variance in the discrete case is Nn_y). In any case we are now ready to establish the CIC formulae analogous to those of Section I. To see the type of argument needed the following result is established (cf. the main result (6e), (7a) of Section Ia).

Proposition 4. For the situation referred to at the beginning of this Section and under the conditions of Proposition 3 (in particular that $\underline{g} \cdot H_0$) we have

$$\hat{\underline{x}}_1(t|T) = \hat{\underline{x}}_0(t|T) + (\underline{c}(t) - \langle \underline{M}_1(t,\cdot),\underline{g}\rangle_1)(\underline{A}\underline{\lambda}_0(0|T) + \underline{m}_1 - \underline{m}_0)$$

where
$$\underline{\lambda}_0(0|T) - \langle \underline{g}',\underline{Y}_0\rangle_0$$

Proof. To simplify the argument and help reveal the basic idea we take $\underline{m}_1 = \underline{m}_0 = 0$

so that $\underline{x}_1(t) = \underline{x}_0(t) = \underline{x}(t)$ and $\underline{y}_1(t) = \underline{y}_0(t) = \underline{y}(t)$ and $\hat{\underline{x}}_i(t|T) = \underline{U}_i(t|T)$

$i = 0, 1.$ Let us denote

$$\hat{\underline{U}}_1(t|T) = \underline{U}_0(t|T) + (\underline{\phi}(t) - \langle \underline{M}_1, \underline{g} \rangle_1) \underline{\Delta V}_0$$

where $$\underline{V}_0 = \underline{\lambda}_0(0|T) - \langle \underline{g}', \underline{y} \rangle_0$$

We have then to show

$$\hat{\underline{U}}_1(t|T) = \underline{U}_1(t|T) - \langle \underline{M}_1, \underline{y} \rangle_1$$

since $\underline{U}_1(t|T)$ is unique this will follow if we establish

$$E_1(\hat{\underline{U}}_1(t|T)\underline{y}'(\sigma)) = \underline{M}_1(t,\sigma) \quad \text{for all} \quad \sigma \quad \text{in} \quad [0,T]$$

Suppose it is shown that

(i) $E_1(\underline{U}_0(t|T)\underline{y}'(\sigma)) = \langle \underline{M}_0(t,\cdot), \underline{S}_1(\cdot,\sigma) \rangle_0$

(ii) $E_1(\underline{V}_0\underline{y}'(\sigma)) = \langle \underline{g}', \underline{S}_1(\cdot,\sigma) \rangle_0$

Then we can find

$$E_1(\hat{\underline{U}}_1(t|T)\underline{y}'(\sigma)) = \langle \underline{M}_0(t,\cdot), \underline{S}_1(\cdot,\sigma) \rangle_0 + (\underline{\phi}(t) - \langle \underline{M}_1(t,\cdot), \underline{g} \rangle_1) \underline{\Delta} \langle \underline{g}', \underline{S}_1(\cdot,\sigma) \rangle_0$$

$$= \langle \underline{M}_0(t,\cdot), \underline{S}_1(\cdot,\sigma) \rangle_0 + \langle \underline{\phi} \underline{\Delta} \underline{g}', \underline{S}_1(\cdot,\sigma) \rangle_0 - \langle \underline{M}_1(t,\cdot), \underline{g} \rangle_1 \underline{\Delta} \langle \underline{g}', \underline{S}_1(\cdot,\sigma) \rangle_0$$

$$= \langle \underline{M}_1(t,\cdot), \underline{S}_1(\cdot,\sigma) \rangle_0 - \langle \underline{M}_1(t,\cdot), \underline{g} \rangle_1 \underline{\Delta} \langle \underline{g}', \underline{S}_1(\cdot,\sigma) \rangle_0 \quad \text{by} \quad (15e)$$

$$= \langle \underline{M}_1(t,\cdot), \underline{S}_1(\cdot,\sigma) \rangle_1 \quad \text{by} \quad (17b)$$

$$= \underline{M}_1(t,\sigma) \quad \text{by the reproducing property}$$

Thus we have to prove (i), (ii).

Let P_1, P_0 be the probability measures induced on the space of sample functions of the normal process $\underline{y}(t)$ under conditions "1", "0" respectively. Now according to Remark 3 following Proposition 2 the assumption $\underline{g} \cdot H_0$ ensures the problem of detection between "1" and "0" is nonsingular so that the Radon-Nikodym derivative

(i.e. likelihood ratio) dP_1/dP_0 exists (cf. Parzen [19]). It follows then that

$$E_1(\underline{U}_0(t|T)\underline{y}'(\sigma)) = E_0(\underline{U}_0(t|T)\underline{y}'(\sigma)dP_1/dP_0)$$

On the other hand

$$E_0(\underline{U}_0(t|T)\underline{y}'(\sigma)) = \underline{H}_0(t,\sigma)$$

by the definition of $\underline{U}_0(t|T)$: However

$$E_0(\underline{y}(t)\underline{y}'(\sigma)dP_1/dP_0) = E_1(\underline{y}(t)\underline{y}'(\sigma))$$

$$= \underline{S}_1(t,\sigma)$$

Now if we can show that

$$\underline{y}(\sigma)dP_1/dP_0 \in SL_2^0 \tag{19}$$

we will have the correspondences between SL_2^0 and H_0

$$\underline{U}_0(t|T) \sim \underline{H}_0(t,\cdot)$$

$$\underline{y}'(\sigma)dP_1/dP_0 \sim \underline{S}_1(\cdot,\sigma)$$

Then it will be possible to apply the definition of $\langle \cdot,\cdot \rangle_0$ to see

$$E_0(\underline{U}_0(t|T)\underline{y}'(\sigma)dP_1/dP_0) = \langle \underline{H}_0(t,\cdot),\underline{S}_1(\cdot,\sigma) \rangle_0$$

which is (i). Similarly (ii) can be established.

To see (19) observe

$$E_1(\underline{y}(\cdot)\underline{y}'(\cdot)) = \underline{S}_1(\cdot,\cdot) \quad H_1$$

So according to (a) of Proposition (3)

$$E_0(\underline{y}(\cdot)\underline{y}'(\cdot)dP_1/dP_0) = \underline{S}_1(\cdot,\cdot) \quad H_0$$

Thus by the definition of H_0 (19) follows. The proof is complete.

Note now that by applying the basic RKHS approximation theorem (Parzen [19]; also Parzen [17, Theorem 6E]; Kailath [5, p. 541]) to the discrete likelihood ratio

expression (14) of Section If we obtain the Radon-Nikodym derivative

$$dP_1/dP_0 = |\underline{A}|\,|\underline{\Sigma}^{-1}|\ \exp\ \{\tfrac{1}{2}\ U\}$$

where

$$U = \frac{1}{2}\ (\hat{\underline{x}}_1(0|T) - \hat{\underline{x}}_0(0|T))'(\underline{P}_1(0,0|T) - \underline{P}_0(0,0|T))^{-1}(\hat{\underline{x}}_1(0|T) - \hat{\underline{x}}_0(0|T))$$

$$- \frac{1}{2}\ (\underline{m}_1 - \underline{m}_0)'\underline{\Delta}^{-1}(\underline{m}_1 - \underline{m}_0)\quad .$$

III. Relation with other work.

Now we demonstrate how these general results include Lainiotis' Generalized Partitioned Algorithm (GPA) (see Govindaraj and Lainiotis, [4]) as well as work of Ljung and Kailath [11].

Consider the state space models

$$\dot{\underline{x}}(t) = \underline{F}(t)\underline{x}(t) + \underline{\Gamma}(t)\underline{w}(t)$$

$$\underline{y}(t) = \underline{H}(t)\underline{x}(t) + \underline{v}(t)$$

or

$$\underline{x}_{k+1} = \underline{F}_k\underline{x}_k + \underline{\Gamma}_k\underline{w}_k$$

$$\underline{y}_k = \underline{H}_k\underline{x}_k + \underline{v}_k$$

where in each case $\underline{w}$, $\underline{v}$ are uncorrelated white noises.

Then each is of the form

$$\underline{x} = \underline{x}_a + \underline{\Phi}\underline{\zeta}$$

$$\underline{y} = \underline{y}_a + \underline{g}\underline{\zeta} = (\underline{H}\underline{x}_a + \underline{v}) + \underline{H}\underline{\Phi}\underline{\zeta}$$

where $\underline{x}_a$, $\underline{y}_a$ are zero-mean processes and $\underline{\zeta} = \underline{x}(0)$. Also in each case $\underline{\Phi}$ is the transition matrix associated with $\underline{F}$. Thus the CIC formula derived above are applicable to these state space models.

IIIa. Relation to GPA.

The partitioning view of the CIC problem begins with condition "0" and imagines condition "1" as produced by adding a further transient to the signal and observation processes. Thus in equation (8a) introduce $\underline{m}_r = \underline{m}_1 - \underline{m}_0$. Also recall (8b) and denote $\underline{\Phi}_r(t,T) = \underline{\Phi}$ to find

$$\underline{\hat{x}}_1(t,T) = \underline{\hat{x}}_0(t,T) + \underline{\Phi}_0(t,0,T)\underline{P}_r(0,T)\underline{\Phi}^{-1}(\underline{m}_r + \underline{\Phi}_0(0,T))$$

Thus we write this as

$$\underline{\hat{x}}_1(t,T) = \underline{\hat{x}}_0(t,T) + \underline{\Phi}_0(t,0,T)\underline{\hat{x}}_r(0,T) \qquad (20a)$$

We rewrite (20a) as

-23-

$$\underline{P}_1(t,t|T) = \underline{P}_0(t,t|T) + \underline{Y}_0(t,0|T)\underline{P}_r(0|T)\underline{Y}_0(t,0|T) \qquad (20b)$$

where we have introduced

$$\hat{\underline{x}}_r(0|T) = \underline{P}_r(0|T)(\underline{\lambda}_0(0|T) + \underline{\Delta}^{-1}\underline{m}_r) \qquad (20c)$$

Finally recall (10c) in the new notation

$$\underline{P}_r^{-1}(0|T) = \underline{Q}_0 + \underline{\Delta}^{-1} \qquad (20d)$$

Now recalling the stochastic interpretation of all the terms in equations (20) set T = t to see these equations can be compared with equations (41), (42), (43), (44) of Govindaraj and Lainiotis [4] (in their notation take $\ell \equiv 0$, $k \equiv t$; "n" $\equiv$ "0" and other obvious equivalences). Also equations (9c), (10f), (11a) can be compared to equations (55), (56), (57).

The case $\underline{m}_0 = \underline{0}$ is discussed, for completeness, in Appendix C.

IIIb. Relation to work of Ljung and Kailath [12].

Equations (9b), (10f) can be compared to Ljung and Kailath [10, equations 32, 33] if we make the equivalence $\underline{W}(t) \equiv \underline{0}$: to see this recall the stochastic interpretation of $\underline{Q}$ in (10d). (There thus appears to be a miswrite of equation (33)). Also equation (8b) can be found in Ljung and Kailath [12] in two cases: for T = t see equation (14); for t = 0 see equation (26) and recall that

$$\underline{Y}_1(0,0|T) = \underline{P}_1(0,0|T)\underline{\Gamma}_1^{-1}$$

$$= \underline{I} - \underline{\Pi}_1\underline{Q}_1 \quad \text{by} \quad (12c) \quad .$$

Conclusion. In this article change of initial condition formulae have been presented

for linear estimates of general nonstationary processes perturbed by fixed rank

transients. The formulae apply to discrete or continuous measurement and to estima-

tion in continuous or discrete time. Also a CIC formula for a Gaussian likelihood

has been given.

The derivations depend essentially on a simple Matrix Inversion Lemma.

Appendix A. A CIC formula for $\underline{P}_1(t,t|T)$.

First observe

$$\hat{\underline{P}}_1(t,t|T) = E_1(\hat{\underline{x}}_1(t|T) - E_1(\underline{x}(t)))[\hat{\underline{x}}_1(t|T) - E_1(\underline{x}(t))]'$$

$$= E_1(\underline{M}_1\underline{S}_1^{-1}\underline{U}_1\underline{U}_1'\underline{S}_1^{-1}\underline{M}_1')$$

$$= \underline{M}_1\underline{S}_1^{-1}\underline{M}_1'$$

The idea is to establish a relation with $\underline{M}_0\underline{S}_0^{-1}\underline{M}_0'$. However let us calculate

$$\underline{P}_1(t,t|T) = E_1(\underline{x}(t) - \hat{\underline{x}}_1(t|T)][\underline{x}'(t) - \hat{\underline{x}}_1(t|T)]$$

$$= E_1(\underline{x}(t) - \hat{\underline{x}}_1(t|T)]\underline{x}'(t) \quad \text{by orthogonality}$$

$$= E_1[\underline{x}(t) - E_1(\underline{x}(t))][\underline{x}'(t)] - E_1[\hat{\underline{x}}_1(t|T) - E_1(\underline{x}(t))]\underline{x}'(t)$$

$$= E[\underline{x}_a(t)\underline{x}_a'(t)] + \underline{\phi}_t\underline{\Pi}_1\underline{\phi}_t' - \hat{\underline{P}}_1(t,t|T) \quad \text{by orthogonality}$$

Thus

$$\underline{P}_1(t,t|T) - \underline{P}_0(t,t|T) = \underline{\phi}_t\underline{\Delta}\underline{\phi}_t' - [\hat{\underline{P}}_1(t,t|T) - \hat{\underline{P}}_0(t,t|T)]$$

Now if it is shown that

$$\hat{\underline{P}}_1(t,t|T) - \hat{P}_0(t,t|T) = \underline{M}_1\underline{S}_1^{-1}\underline{M}_1' - \underline{M}_0\underline{S}_0^{-1}\underline{M}_0'$$

$$= -(\underline{\phi}_t - \underline{M}_1\underline{S}_1^{-1}\underline{G})\underline{\Delta}(\underline{\phi}_t - \underline{G}'\underline{S}_0^{-1}\underline{M}_0') + \underline{\phi}_t\underline{\Delta}\underline{\phi}_t'$$

then the result (12a) will follow in view of (7c).

Consider then

$$\underline{M}_1\underline{S}_1^{-1}\underline{M}_1' = \underline{M}_1(\underline{S}_0^{-1} - \underline{S}_1^{-1}\underline{G}\underline{\Delta}\underline{G}'\underline{S}_0^{-1})\underline{M}_1' \quad \text{by} \quad \text{(MILb)}$$

$$= \underline{M}_1\underline{S}_0^{-1}\underline{M}_1' - \underline{M}_1\underline{S}_1^{-1}\underline{G}\underline{\Delta}\underline{G}'\underline{S}_0^{-1}\underline{M}_1'$$

$$= (\underline{M}_0 + \underline{\phi}_t\underline{\Delta}\underline{G}')\underline{S}_0^{-1}\underline{M}_1' - \underline{M}_1\underline{S}_1^{-1}\underline{G}\underline{\Delta}\underline{G}'\underline{S}_0^{-1}\underline{M}_1' \quad \text{by} \quad \text{(6c)}$$

$$= \underline{M}_0\underline{S}_0^{-1}\underline{M}_1' + (\underline{\phi}_t - \underline{M}_1\underline{S}_1^{-1}\underline{G})\underline{\Delta}\underline{G}'\underline{S}_0^{-1}\underline{M}_1'$$

$$= M_0 S_0^{-1} M_0' + M_0 S_0^{-1} G \Delta \phi_t' + (\phi_t - M_1 S_1^{-1} G) \Delta G' S_0^{-1} M_1'$$

$$= M_0 S_0^{-1} M_0' + M_0 S_0^{-1} G \Delta \phi_t' + (\phi_t - M_1 S_1^{-1} G) \Delta G' S_0^{-1} G \Delta \phi_t' + (\phi_t - M_1 S_1^{-1} G) \Delta G' S_0^{-1} M_0'$$

Now recall (9a)

$$\phi_t - M_1 S_1^{-1} G = (\phi_t - M_0 S_0^{-1} G)(\underline{I} - \Delta G' S_1^{-1} G)$$

which, by (10b) is also

$$= (\phi_t - M_0 S_0^{-1} G)(I + \Delta G' S_0^{-1} G)^{-1}$$

Thus the sum of the <u>second</u> and <u>third</u> terms above is

$$M_0 S_0^{-1} G \Delta \phi_t' + (\phi_t - M_0 S_0^{-1} G)(\underline{I} + \Delta G' S_0^{-1} G)^{-1}(\Delta G' S_0^{-1} G) \Delta \phi_t'$$

$$= M_0 S_0^{-1} G \Delta \phi_t' + (\phi_t - M_0 S_0^{-1} G)(I - (I + \Delta G' S_0^{-1} G)^{-1}) \Delta \phi_t'$$

$$= M_0 S_0^{-1} G \Delta \phi_t' + (\phi_t - M_0 S_0^{-1} G) \Delta \phi_t' - (\phi_t - M_1 S_1^{-1} G) \Delta \phi_t'$$

$$= \phi_t \Delta \phi_t' - (\phi_t - M_1 S_1^{-1} G) \Delta \phi_t'$$

Thus

$$M_1 S_1^{-1} M_1' = M_0 S_0^{-1} M_0' + \phi_t \Delta \phi_t' - (\phi_t - M_1 S_1^{-1} G) \Delta \phi_t' + (\phi_t - M_1 S_1^{-1} G) \Delta G' S_0^{-1} M_0'$$

$$= M_0 S_0^{-1} M_0 + \phi_t \Delta \phi_t' - (\phi_t - M_1 S_1^{-1} G) \Delta (\phi_t - G' S_0^{-1} M_0')$$

which is what was required to be established.

<u>Appendix B</u>. Derivation of (16).

Let $\underline{a}_i$ $i = 1, \ldots, N$ be any m-vectors and $t_1 < t_2 < \ldots < t_N$ be any N points in $[0,T]$. The claim is that

$$E = \sum_{i=1}^{N} \sum_{j=1}^{N} \underline{a}_i' \underline{\Gamma}(t_i) \underline{D} \underline{\Gamma}'(t_j) \underline{a}_j \leq \lambda_D \operatorname{tr}(\langle \underline{\Gamma}', \underline{\Gamma} \rangle_\Sigma) \sum_{i=1}^{N} \sum_{j=1}^{N} \underline{a}_i' \underline{\Sigma}(t_i, t_j) \underline{a}_j$$

Now

$$E = (\sum_{i=1}^{N} \underline{a}_i' \underline{\Gamma}(t_i)) \underline{D} (\sum_{j=1}^{N} \underline{\Gamma}'(t_j) \underline{a}_j)$$

$$= \underline{u}' \underline{D} \underline{u} \leq \lambda_D \underline{u}' \underline{u}$$

where

$$\underline{u} = \sum_{i=1}^{N} \underline{\Gamma}'(t_i) \underline{a}_i$$

Next apply the reproducing property to see

$$\underline{u} = \sum_{i=1}^{N} \langle \underline{\Gamma}'(\cdot), \underline{\Sigma}(\cdot, t_i) \rangle_\Sigma \underline{a}_i$$

$$= \langle \underline{\Gamma}'(\cdot), \sum_{i=1}^{N} \underline{\Sigma}(\cdot, t_i) \underline{a}_i \rangle_\Sigma$$

$$= \langle \underline{\Gamma}'(\cdot), \underline{w}_N(\cdot) \rangle_\Sigma \qquad \text{say}$$

Now $\underline{u}$ has kth element u_k where

$$u_k = \langle \underline{\gamma}_k', \underline{w}_N \rangle_\Sigma : \quad \underline{\gamma}_k \text{ is the kth column of } \underline{\Gamma}$$

Thus

$$\underline{u}' \underline{u} = \sum_{k=1}^{d} u_k^2 \leq \sum_{k=1}^{d} \langle \underline{\gamma}_k', \underline{\gamma}_k \rangle_\Sigma \langle \underline{w}_N', \underline{w}_N \rangle_\Sigma$$

$$= \operatorname{tr} \langle \underline{\Gamma}', \underline{\Gamma} \rangle_\Sigma \langle \underline{w}_N', \underline{w}_N \rangle$$

However

$$\langle \underline{w}_N', \underline{w}_N \rangle_\Sigma = \langle \sum_{i=1}^{N} \underline{a}_i' \underline{\Sigma}(\cdot, t_i), \sum_{i=1}^{N} \underline{\Sigma}(\cdot, t_i) \underline{a}_i \rangle_\Sigma$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{N} \underline{a}_i' \langle \underline{\Sigma}(\cdot, t_i), \underline{\Sigma}(\cdot, t_j) \rangle_\Sigma \underline{a}_j$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{N} \underline{a}_i' \underline{\Sigma}(t_i, t_j) \underline{a}_j$$

by the reproducing property. The result is thus established.

-28-

<u>Appendix C</u>. Relation with partitioned algorithm for $\underline{\Xi}_0 = 0$.

Consider now the case that $\underline{\Xi}_0 = 0$. Then by (10a)

$$P_r(0|T) = \underline{\Delta} - \underline{\Delta}\underline{C}_1\underline{\Delta}$$

$$= \underline{\Xi}_1 - \underline{\Xi}_1\underline{C}_1\underline{\Xi}_1 = P_1(0,0|T) = \underline{P}_1$$

Thus, since $\underline{P}_r = \underline{V}$, (9b) becomes

$$\underline{\Phi}_1(t,0|T)\underline{V}_1 = \underline{\Phi}_0(t,0|T)\underline{P}_1 \tag{C1}$$

while (9d) reads

$$\underline{\Phi}_1(t,0|T) = \underline{\Phi}_0(t,0|T) - \underline{\Phi}_0(t,0|T)\underline{C}_0\underline{P}_1 \tag{C2}$$

Also (10c) becomes

$$\underline{P}_1^{-1} = \underline{\Xi}_1^{-1} + \underline{C}_0 \tag{C3}$$

When further $\underline{m}_0 = \underline{0}$, (7b) implies

$$\hat{\underline{x}}_0(0|T) = \underline{0}$$

Thus for $t = 0$ (8a) becomes

$$\hat{\underline{x}}_1(0|T) = \underline{P}_1(\underline{\Xi}_1^{-1}\underline{m}_1 + \underline{\lambda}_0(0|T)) \tag{C4}$$

The pair (C3), (C4) with $T = t$, are just a general version of Lainiotis' original partitioned algorithm: Lainiotis [9], equations (3), (4). Equations (C1), (C2) can be compared with Lainiotis [9] equations (20), (21).

-29-

REFERENCES

[1] H. B. Aasnaes and T. Kailath, "Initial-Condition Robustness of Linear Least
     Squares Filtering Algorithms," IEEE Trans. Autom. Contr., VolAC-19, pp. 393-
     397, 1974.

[2] N. Aronszajn, "Theory of Reproducing Kernels," Trans. Amer. Math. Soc., vol. 68,
     pp. 337-404, 1950.

[3] D. G. Fraser, "A New Technique for Optimal Smoothing of Data," Sc.D. Dissertation,
     M.I.T., 1967.

[4] K. S. Govindaraj and D. G. Lainiotis, "A Unifying Framework for Discrete Linear
     Estimation, Generalised Partitioned Algorithms," Int. J. Control., vol. 28,
     pp. 571-588, 1978.

[5] T. Kailath, "RKHS Approach to Detection and Estimation Problems - Part I:
     Deterministic Signals in Gaussian Noise," IEEE Trans. Inf. Thy., vol. IT-17,
     pp. 530-549, 1971.

[6] T. Kailath, "RKHS Approach to Detection and Estimation Problems - Part II:
     Gaussian Signal Detection," IEEE Trans. Inf. Thy., vol. IT-21, pp. 15-23, 1975.

[7] T. Kailath, R. T. Geesey and H. L. Weinert," Some Relations Among RKHS Norms,
     Fredholm Equations, and Innovations Representations," IEEE Trans. Inf. Thy.,
     vol. IT-18, pp. 341-348, 1972.

[8] T. Kailath and P. Frost, "An Innovations Approach to Least-Squares Estimation
     Part II: Linear Smoothing in Additive White Noise, IEEE Trans. Autom. Contr.,
     vol. AC-13, pp. 655-660, 1968.

[9] D. G. Lainiotis, "Partitioned Estimation Algorithms, II: Linear Estimation,"
     Inf. Sci., vol 7, pp. 317-340, 1974.

[10] D. G. Lainiotis, "A Unifying Framework for Linear Estimation: Generalised
     Partitioned Algorithms," Inf. Sci., vol. 10, pp. 243-278, 1976.

[11]   L. Ljung and T. Kailath, "A Unified Approach to Smoothing Formulas," Automatica, vol. 12, pp. 147-157, 1976.

[12]   L. Ljung and T. Kailath, "Efficient Change of Initial Conditions, Dual Chandrasekhar Equations; and Some Applications," IEEE Trans. Autom. Contr., vol. AC-22, pp. 443-447, 1977.

[13]   D. Q. Mayne, "A Solution of the Smoothing Problem for Linear Dynamic Systems," Automatica, vol. 4, pp. 73-92, 1966.

[14]   R. K. Mehra, "On Optimal and Suboptimal Linear Smoothing," Tech. Rep. No. 559, Div. Eng. Appl. Phys., Harvard Univ., 1968.

[15]   H. Oodaira, "Equivalence and Singularity of Gaussian Measures," Ph.D. Thesis, Dept. Math., Mich. State Univ., E. Lansing Mich., 1963.

[16]   E. Parzen, "Statistical Inference on Time Series by Hilbert Space Methods," in "Time Series Analysis Papers" by E. Parzen, Holden-Day, 1967.

[17]   E. Parzen, "A New Approach to the Synthesis of Optimal Smoothing and Prediction Systems," in "Time Series Analysis Papers" by E. Parzen, Holden-Day, 1967.

[18]   E. Parzen, "Extraction and Detection Problems and Reproducing Kernel Hilbert Spaces," SIAM. J. Contr., ser. A., vol. 1, pp. 492-519, 1962.

[19]   E. Parzen, "Probability Density Functionals and Reproducing Kernel Hilbert Spaces," in Proc. of Symposium of Time Series, M. Rosenblatt ed. Brown Univ., 1962.

[20]   E. Parzen, "An Approach to Time Series Analysis," Annals. Math. Stat., vol. 32, pp. 951-989, 1961. Also reprinted in "Time Series Analysis Papers" by E. Parzen, Holden-Day, 1967.

[21]   H. E. Rauch, F. Tung and C. T. Striebel, "Maximum Likelihood Estimates of Linear Dynamic Systems," AIAA J., vol. 3, pp. 1445-1450, 1965.

VS/ck

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER 2119	2. GOVT ACCESSION NO. AD-A093633	3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle) Smoothing Estimation of Stochastic Processes Part I. Change of Initial Condition Formulae.	5. TYPE OF REPORT & PERIOD COVERED Summary Report, no specific reporting period
	6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s) V. Solo	8. CONTRACT OR GRANT NUMBER(s) DAAG29-80-C-0041

9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of 610 Walnut Street Wisconsin Madison, Wisconsin 53706	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 4 – Statistics and Probability

11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P. O. Box 12211 Research Triangle Park, North Carolina 27709	12. REPORT DATE September 1980
	13. NUMBER OF PAGES 31

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report) UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Linear Estimation, Smoothing, Filtering, Covariance Kernel, Hilbert Space

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The change of initial condition (CIC) problem in the theory of smoothing in linear estimation is formulated in terms of a fixed rank perturbation to a covariance kernel. The CIC or partitioning formulae are then shown to apply for general nonstationary processes. It becomes clear that the formulae involved derive from inversion formulae for fixed rank modification of a positive definite kernel or matrix.

DD 1 JAN 73 1473   EDITION OF 1 NOV 65 IS OBSOLETE

DAT
FILM